

AD _____

GRANT NUMBER DAMD17-96-1-6228

TITLE: Demonstration Project on Mammographic Computer-Aided Diagnosis for Breast Cancer Detection

PRINCIPAL INVESTIGATOR: Kunio Doi, Ph.D.

CONTRACTING ORGANIZATION: University of Chicago
Chicago, IL 60637

REPORT DATE: October 1997

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

DTIC QUALITY INSPECTED 2

19980130 181

REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY <i>(Leave blank)</i>	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED	
	October 1997	Annual (26 Sep 96 - 25 Sep 97)	
4. TITLE AND SUBTITLE Demonstration Project on Mammographic Computer-Aided Diagnosis for Breast Cancer Detection			5. FUNDING NUMBERS DAMD17-96-1-6228
6. AUTHOR(S) Kunio Doi, Ph.D.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Chicago Chicago, IL 60637			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command Fort Detrick, Frederick, Maryland 21702-5012			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited		12b. DISTRIBUTION CODE	
13. ABSTRACT <i>(Maximum 200)</i> The goal of this project is to demonstrate the clinical usefulness of computer-aided diagnosis (CAD) in mammographic detection of breast cancer. Our plan is to develop advanced CAD schemes for detection and characterization of clustered microcalcifications and masses by incorporating artificial neural networks and various image processing techniques. Clinical mammography workstations for automated detection of suspicious lesions in mammograms will be developed by integration of laser digitizer, high-speed computer and advanced CAD software. The prototype workstations will be used as a "second opinion" in interpreting mammograms by reducing observational errors. The outcomes of radiologists' image readings in the detection of breast cancer will be evaluated by examining radiologists' performance when reading films only and when reading film with the computer results. We believe that the outcomes of this demonstration project will lead to large-scale clinical trials and will result in commercial products for practical routine use in breast imaging.			
14. SUBJECT TERMS Breast Cancer			15. NUMBER OF PAGES 31
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

KD Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

KD For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Kunio Dai 10/2/97
PI - Signature Date

Table of Contents

	page
1. INTRODUCTION.....	2
1.1 Subject and Scope of the Research	2
1.2 Purpose	5
1.3 Background of Previous Work.....	6
2. BODY	
2.1 Experimental Methods, Assumptions and Procedures	11
2.2 Results and Discussions	16
(1) Development of automated detection scheme for	16
clustered microcalcifications	
(2) Development of automated detection scheme for masses ...	17
(3) Development of automated scheme for characterization	
of clustered microcalcifications.....	18
(4) Development of automated scheme for characterization	
of masses	19
(5) Development of prototype CAD workstations.....	20
(6) Clinical evaluation of CAD workstations	21
2.3 Recommendations in relation to the Statement of Work.....	23
3. CONCLUSIONS	24
4. REFERENCES	24

1. INTRODUCTION

1.1. The Subject and Scope of the Research

Breast cancer is a leading cause of death in women, with an estimated 46,000 deaths per year in the United States (ref. 1). Mammography is currently the only known reliable method for early detection of breast cancer (refs. 2,3). However, early mammographic signs of breast cancer such as clustered microcalcifications and masses are usually very subtle, and thus 10-30% of lesions are missed even by trained radiologists. These misses are due to the often low conspicuity of lesions, eye fatigue, and human error (refs. 4-7). However, there is clear evidence (refs. 8,9) that radiologists' accuracy in the detection of subtle breast lesions would be improved if a computer output indicating possible sites of suspicious lesions were made available to radiologists as a "second opinion."

As a team of investigators at the Kurt Rossmann Laboratories for Radiologic Image Research at the University of Chicago, we have been involved since 1985 in developing the concepts and methodology of computer-aided diagnosis (CAD) with which to assist radiologists in detecting lesions and improving the sensitivity of breast cancer diagnosis through mammography (refs. 10, 11). CAD may be defined as a diagnosis made by a radiologist who takes into account the results of automated computer analyses of radiographic images. The computer output may be used as a "second opinion." We have extensive experience in developing CAD schemes. In addition to breast cancer, we have developed computer schemes for the detection of lung nodules (refs. 12, 13), interstitial infiltrates (refs. 14, 15), cardiomegaly (refs. 16, 17) and pneumothorax (ref. 18) in chest radiography; the detection of stenotic lesions and blood flow analysis in angiography (refs. 19, 20); and the assessment of osteoporosis in skeletal radiography (ref. 21).

In mammography, CAD schemes are being developed for detection of clustered microcalcifications (refs. 8, 24-26, 28) and for detection of masses (refs. 22, 23). A basic scheme for automated detection of clustered microcalcifications employs a difference image

technique to enhance the signal-to-noise ratio of microcalcifications, followed by thresholding, feature extraction and classification using artificial neural networks. At present, the performance of this scheme provides a sensitivity of approximately 85% in the detection of clustered microcalcifications with a false positive rate of approximately 0.7 per mammogram when it is tested on our database of 78 mammograms, in which one half are normal cases and the other half includes subtle clustered microcalcifications.

For the automated detection of mammographic masses, another CAD scheme is being developed on the basis of a bilateral subtraction technique that analyzes deviations of architectural symmetry between the right and left breast images, with asymmetries indicating potential masses (refs. 29-32). Currently, this scheme performs at approximately 90% sensitivity with a false positive rate of about 2 per case when it is tested on our database of 154 pairs of mammograms. Our current research effort on these CAD schemes is focused primarily on improving further their performance through careful analysis of computer false-positives and false-negatives.

To date, these studies have been performed retrospectively on selected sets of mammograms, and we have obtained results that indicate that our schemes have the potential to be used as an effective aid for radiologists. We are now at the stage in the development of our CAD program to test our schemes prospectively on a large number of clinical mammograms.

On November 8th, 1994, we implemented an "intelligent" mammography workstation (ref. 34) and began the first test of our schemes on clinical screening mammograms obtained in the mammography section of our department. This workstation consists of an IBM Powerstation 590, a Konica LD4500 laser film digitizer, an Alphatronix Inspire 40-GB magneto-optical jukebox, two Imlogix 1024 line monitors, and a Seikosha VP 4500 video printer for hard copy. The "intelligence" of the workstation comes from our automated detection schemes for clustered microcalcifications and masses.

In order to realize clinically and practically mammographic CAD for detection of breast cancers in screening programs, it is necessary to have commercial products for widespread use by radiologists in breast clinics, community hospitals, and academic medical centers. Therefore, in 1993, ARCH Development Corporation (ARCH), which is a not-for-profit organization created by the Board of Trustees of the University of Chicago in 1986 as a unique mechanism to commercialize inventions developed by the faculty at the University of Chicago and by scientists at Argonne National Laboratory, licensed its inventions on CAD and related technologies to R2 Technology, Inc.

R2 Technology, Inc. was founded in 1993 with the specific goal of developing and marketing a computer aided diagnostic system in mammographic detection of breast cancer. R2 Technology, Inc., has been funded by leading venture firms in Silicon Valley--Sigma Partners, and Burr, Egan, Deleage and Co. -- that have supported many other successful medical and computer companies. Its development group has over 250 man-years of experience in medical imaging and computer systems. Its product development process has been to identify those high potential prototype systems in leading research institutions and form alliances and integrate those systems into R2's core technology. As of May 1995, R2 has alliances with the University of Chicago, Lockheed Missiles & Space Company, Inc., and Sandia National Laboratories.

Therefore, the next logical step in the development of CAD is to conduct a large-scale, multi-institutional demonstration project to examine whether additional breast cancers can be found by use of mammographic CAD workstations. We believe that the performance of mammographic CAD schemes has reached the high level necessary for clinical evaluation. Serious efforts toward commercialization of CAD units have already begun. Therefore, it is likely that a clear positive outcome from this study would result in production of commercial products for widespread use in breast imaging and would lead to large-scale clinical trials.

1.2. Purpose

The goal of this project is to demonstrate the clinical usefulness of computer-aided diagnosis (CAD) in mammographic detection of breast cancer through multi-disciplinary and multi-institutional efforts. We plan to develop clinical prototype mammography workstations for automated detection of suspicious lesions in mammograms by incorporating image processing techniques and artificial neural networks. The prototype workstation will be used as a "second opinion" to assist radiologists' interpretation of mammograms. Clinical usefulness of the mammography CAD will be demonstrated and evaluated at four hospitals in the Chicago area. The major hypothesis to be tested in this proposal is that CAD improves accuracy in the detection of breast cancer by reducing observational errors on mammographic images. Our proposal is designed to demonstrate that approximately 23 additional breast cancers will be detected among approximately 45,000 screenees due to the use of CAD computer output.

The specific aims of this demonstration project are listed below.

- (1) Further development of advanced CAD schemes for detection of breast lesions
 - (a) Automated detection scheme for clustered microcalcifications
 - (b) Automated detection scheme for masses
 - (c) Automated scheme for characterization of detected breast lesions
- (2) Development of the prototype mammography CAD workstations by integration of laser digitizer, high-speed computer, and advanced CAD schemes
- (3) Clinical demonstration and evaluation of prototype mammography CAD workstations at two hospitals: one academic institution and one community hospital
- (4) Analysis of outcomes of the clinical evaluation of the prototype workstations for detection of additional breast cancers by the use of computer output

1.3. Background of Previous Work

We have been working on the development of computer-aided diagnostic (CAD) schemes for mammography, chest radiography, angiography, and bone radiography since 1985. Therefore, we have extensive experience in quantitative analysis of radiographic images for detection and characterization of various patterns based on computer-vision methods and artificial neural networks. These extensive studies provide the basis for the continued development and testing of advanced CAD schemes for the detection of breast lesions proposed in this research. A number of investigations which are relevant to this study are described briefly here.

(1) Development of computerized detection scheme for mammographic microcalcifications

We have investigated the application of computer-based methods to the detection of microcalcifications on digital mammograms. Our computer detection system was based on a difference-image technique in which a signal-suppressed image was subtracted from a signal-enhanced image to remove the structured background in a mammogram (ref. 24). Signal-extraction techniques adapted to the known physical characteristics of microcalcifications were then used to isolate microcalcifications from the remaining noise background (ref. 25). Signal-extraction criteria based on the size, contrast, number, texture, and clustering properties of microcalcifications were next imposed on the detected signals to distinguish true signals from noise or artifacts (refs. 8, 25). The detection accuracy of the computer scheme was evaluated by means of a free-response receiver operating characteristic (FROC) analysis. In a study of 78 clinical images containing subtle microcalcifications, the automated computer scheme obtained an 85% true-positive cluster detection rate at a false-positive detection rate of 1.5 clusters per image. These results indicated that the automated method has the potential to aid radiologists in screening mammograms for clustered microcalcifications.

We have applied a shift-invariant neural network (SIANN) to eliminate false-positive detections reported by the CAD scheme. The SIANN is a layered feed-forward

neural network with local, spatially-invariant interconnections (refs. 27, 28). The basic idea of local, spatially-invariant interconnections (or sharing local interconnection weights) was first introduced by Fukushima in his Neocognitron for recognition of handwriting characters in the early 1980s (refs. 35, 36). The SIANN developed by Zhang et al. (ref. 27) for image processing is a feed-forward neural network without the lateral interconnections and feedback loops that are included in the Neocognitron. Furthermore, a modified error backpropagation (EBP) algorithm with the shift-invariant-connection constraint (ref. 27) is used as the training algorithm in the SIANN. The SIANN has been shown to be a powerful tool for pattern recognition and image processing, since it can learn to discriminate between objects on the basis of local features with results that are invariant to translation of the objects (refs. 27, 28).

This neural network was trained to detect each individual microcalcification in a given region of interest (ROI) reported by the CAD scheme. A ROI was classified as a positive ROI if the total number of microcalcifications detected in the ROI was greater than two. The performance of the shift-invariant neural network was evaluated by means of a jack-knife method and conventional receiver operating characteristic (ROC) analysis by using a database of 168 ROIs that had been reported by the CAD scheme when applied to 39 mammograms. The analysis yielded an average area under the ROC curve (A_z) of 0.91. Approximately 55% of false-positive ROIs were eliminated without any loss of true-positive ROIs (ref. 28). This result was considerably better than that obtained in our previous study using a conventional three-layer, feed-forward neural network.

We have also studied radiologists' detection of clustered microcalcifications on mammograms to determine whether CAD can improve radiologists' performance. The results of a ROC study showed that CAD, using the level of computer performance at that time (sensitivity = 87%, 4 false clusters per image), does significantly ($p<0.001$) improve radiologists' accuracy in detecting clustered microcalcifications under conditions that simulate the rapid interpretation of screening mammograms (ref. 8). The results also

suggested that a reduction in the computer's false-positive rate would further improve radiologists' diagnostic accuracy.

The importance of our findings is that a computerized scheme can detect clustered microcalcifications in digitized mammograms at a high level of sensitivity that would be comparable to levels obtained by radiologists. In addition, radiologists' performance in the detection of clustered microcalcifications can be improved significantly when the results of the computer output are provided as an aid to the radiologists.

(2) Development of computerized detection schemes for mammographic masses

A computerized scheme has been developed for the detection of masses in digital mammograms. Based on deviations from the normal architectural symmetry of the right and left breasts, a bilateral subtraction technique was used to enhance the conspicuity of possible masses. The scheme employed pairs of conventional screen-film mammograms (right and left MLO views and right and left CC views), which were digitized by a TV camera/Gould digitizer. The right and left breast images in each pair were aligned manually during digitization. A nonlinear bilateral subtraction technique, which involves linking multiple subtracted images, was investigated and compared to a simple linear subtraction method (refs. 29, 30). Various feature-extraction techniques were used to reduce false-positive detections resulting from the bilateral subtraction. The scheme was evaluated using 46 pairs of clinical mammograms and was found to yield a 95% true-positive rate at an average of three false-positive detections per image. This preliminary study indicated that the scheme is potentially useful as an aid to radiologists in the interpretation of screening mammograms.

We continued to investigate the characteristics of actual masses and non-mass detections in order to develop feature-analysis techniques with which to reduce the number of non-mass (i.e., false-positive) detections. These feature-analysis techniques involved extraction of various features such as area, contrast, circularity and border-distance based

on the density and geometric information of masses in both processed and original breast images. Cumulative histograms of both actual-mass detections and non-mass detections were used to characterize extracted features and to determine the cutoff values used in the feature tests. The effectiveness of the feature-analysis techniques was evaluated using FROC analysis. Results showed that the feature-analysis techniques effectively improved the performance of the computerized detection scheme: about 35% of false-positive detections were eliminated without loss in sensitivity (ref. 31).

We have developed an automated technique for the alignment of right and left breast images for use in the computerized analysis of bilateral breast images. In this technique (ref. 32), the breast region was first identified by use of histogram analysis and morphological operations. The anterior portions of the tracked breast border and computer-identified nipple positions were selected as landmarks for image registration. The paired right and left breast images were then registered relative to each other by use of a least-squares matching method. Based on FROC and regression analyses, the detection performance obtained with the automated alignment technique was found to be higher than that obtained with simulated misalignments. These results indicated that automatic alignment of breast images is feasible and that mass-detection performance appears to improve with the inclusion of asymmetric anatomic information and is not sensitive to slight misalignment.

We also investigated the effect of case selection on the performance of a CAD scheme, since the choice of clinical cases used to test the scheme can affect the test results. In this study, we deliberately modified the components of our database to study the effects of this modification on measured performance. Using our computerized scheme for the automated detection of breast masses from mammograms, we found that the sensitivity of the scheme ranged between 26% to 100% (at a false positive rate of 1.0 per image), depending on the cases used to test the scheme. Even a 20% change in the cases comprising the database reduced the measured sensitivity by 15-25% (ref. 33). Because of

the strong dependence of measured performance on the testing database, it is difficult to estimate reliably the accuracy of a CAD scheme. Moreover, it is questionable to compare different CAD schemes when different cases are used for testing. Sharing databases, creating a common database, or using a quantitative measure to characterize databases are possible solutions to this problem. However, none of these solutions exists or is practiced at present. Therefore, as a short-term solution, we recommend that the method used for selecting cases and histograms of relevant image features be reported whenever performance data are presented.

The importance of our findings is that a nonlinear bilateral subtraction technique can detect mammographic masses at a high level of sensitivity that are again comparable to levels obtained by radiologists.

(3) Computed Detection of Lesions Missed by Mammography

Over the past 6 years, we have been collecting cases in which a lesion was missed in a mammogram. To date, 69 cases with a lesion that went undetected by a radiologists were analyzed by the two detection schemes -- clustered microcalcifications and masses (ref. 37). In all cases the lesions were rated retrospectively as being subtle to extremely subtle by an experienced radiologist. The computer schemes correctly identified approximately 50% of the missed lesions -- 54% of the malignant lesions and 45% of the benign lesions. The false positive rate was 1.3 per image. This result shows that our computer detection schemes are capable of identifying cancers that are overlooked by radiologists.

(4) Classification Schemes

We have developed a method for differentiating malignant from benign clustered microcalcifications in which image features are both extracted and analyzed by a computer. One hundred mammograms obtained from 53 patients who had biopsies for suspicious

clustered microcalcifications were used. Our technique used 8 computer-extracted features of clustered microcalcifications that were merged by an artificial neural network. Features were based on the size and shape of clusters and on the size, shape, contrast, and uniformity of individual microcalcifications comprising a cluster. Human input was limited to initial identification of the microcalcifications. Our method correctly classified 100% of patients with breast cancer and 69% of patients with biopsy-proven benign conditions. ROC analysis showed that our method performed significantly ($p=0.03$) higher than 5 radiologists who reviewed the mammograms retrospectively. This result indicated that quantitative features extracted by a computer can be analyzed by a computer to distinguish malignant from benign clustered microcalcifications, and that our technique can potentially help radiologists to reduce the number of "false-positive" biopsies.

2. BODY

2.1. Experimental Methods, Assumptions and Procedures

The overall plan of this demonstration project involves four major steps, namely, (1) further development of advanced CAD schemes, (2) development of prototype mammography CAD workstations, (3) clinical evaluation of prototype workstations, and (4) analysis of outcomes from clinical evaluations.

The primary goal of this study is to demonstrate that approximately 23 additional breast cancers will be detected by the use of prototype mammography CAD workstations for approximately 45,000 screenees who are expected to enter a three-year clinical evaluation at two hospitals. The potential of detecting 23 additional breast cancers was estimated from an average breast cancer incidence rate of five per 1,000 screenees, a current average miss rate of 20%, and a level of CAD performance that detects 50% of currently missed cancer lesions.

Advanced CAD schemes will be developed for detection of clustered microcalcifications and masses as well as characterization of detected lesions by integrating

a number of new methods into the existing programs and optimizing a number of parameters for achieving high performance levels above the current ones. Two kinds of prototype mammography CAD workstations will be developed. The first prototype unit is based on the existing intelligent workstation at the University of Chicago, which will incorporate the most advanced CAD software and will be used for clinical evaluation on approximately 30 screenees per day at the University of Chicago. The second type is the prototype commercial units which will be developed by R2 Technology, Inc., and will be used for clinical evaluation on approximately 30 screenees per day at LaGrange Memorial Hospital.

The impact of the computer output from the prototype workstation will be evaluated by examining if and when the radiologist changes his/her initial diagnosis. The computer output will be presented to the radiologist only after the radiologist has entered his/her initial findings into the computer as to the normal and abnormal lesion(s). A particularly important datum in this demonstration project is the measurement of the number of breast cancer cases on which the radiologist did not initially indicate the breast cancer lesion but did make a final correct diagnosis by using the computer output as a "second opinion."

In this demonstration project, we will not direct effort toward the development of major new methods and techniques on mammographic CAD schemes. Instead, we plan to incorporate several useful methods and techniques, which are recently developed, into the CAD software package for implementation in the prototype intelligent mammography workstation. It is important to note that considerable research effort would be required to optimize many parameters associated with new CAD methods and the existing CAD algorithms in order to integrate all of the components into a single package that functions successfully.

In the first phase of this project, we plan to develop advanced CAD schemes for detection of clustered microcalcifications and masses, and then to incorporate them into the prototype intelligent mammography workstation for clinical evaluation at the University of

Chicago. However, as the performance of advanced CAD schemes in our laboratory improves through continued efforts on the optimization process, the CAD software package in the workstation will be upgraded as needed. In the second phase of this project, we plan to incorporate additional CAD schemes to characterize detected lesions as benign or malignant.

(1) Automated scheme for detection of clustered microcalcifications

We plan to investigate and incorporate three new approaches to improve the performance of automated detection of clustered microcalcifications. They are (1) local edge-gradient analysis techniques for reduction of false-positives, (2) shift-invariant neural networks for removal of false-positives, and (3) wavelet transform techniques for improvement in the sensitivity in detecting clustered microcalcifications, as described below. Many parameters associated with these approaches will be selected carefully to optimize the overall performance in detecting clustered microcalcifications. It is important to note that previous studies on these methods were based on mammograms digitized using a drum scanner. In this project, we plan to determine all of the new parameters with mammograms digitized using a laser scanner that is integrated into the prototype intelligent mammography workstation.

(2) Automated scheme for detection of masses

We plan to investigate and incorporate three new approaches to improve the performance of automated detection of mass lesions: (1) Hough spectrum analysis for the detection of spiculated lesions and architectural distortions; (2) gradient and circularity analysis for the detection of very small early cancers; and (3) artificial neural networks for the merging of various features of suspect lesions, identified either by the bilateral subtraction method or by the two new single image methods, in order to reduce the number of false-positive detections.

(3) Automated scheme for characterization of detected lesions

In the second phase of development of advanced CAD schemes, we plan to investigate and incorporate two automated schemes for distinguishing between benign and malignant lesions both for detected clustered microcalcifications and masses. The likelihood of malignancy on each detected suspicious lesions will be calculated from our schemes and will be displayed together with the location(s) of detected lesion(s) on the prototype mammography CAD workstation at the University of Chicago. We plan to investigate whether the calculated likelihood of malignancy added to the CAD computer output may improve the diagnosis of breast cancer by reducing the false-positives and false-negatives.

(4) Development of prototype mammography CAD workstations

We plan to develop two kinds of prototype mammography CAD workstations for clinical evaluation. One is based on the existing intelligent mammography workstation at the University of Chicago, which will be modified by incorporating advanced CAD software and by improving some aspects of the hardware configuration. This first prototype system will be used for clinical evaluation at the University of Chicago. The second type of prototype system will be developed by R2 Technology, Inc., as a potential commercial unit, and will be placed for clinical evaluation at LaGrange Memorial Hospital. Although the basic principles employed in the two kinds of prototype workstations are similar due to licensing of the University of Chicago technologies to R2 Technology, Inc., these two systems are not identical. Therefore, we plan to investigate the levels of performance of each prototype workstation.

(5) Clinical evaluation of prototype mammography CAD workstations

Multi-institutional clinical evaluation of mammography CAD workstations will be carried out at two clinical sites: the Mammography Section of the Department of Radiology,

the University of Chicago and LaGrange Memorial Hospital in LaGrange, Illinois. The number of screenees per day who will enter this clinical evaluation at each of the two hospitals is approximately 30. The total number of screenees per day will be 60. We have already obtained an approval from the Institutional Review Board (IRB) for clinical evaluation of the prototype intelligent mammography workstation at the University of Chicago and LaGrange Memorial Hospital.

To examine the impact of mammographic CAD on clinical outcomes, we plan to obtain data from mammography audits without and with the prototype CAD workstations. For the first six months of this project, the CAD workstation will not be used and we will collect results of mammography audits. For the next year, the first clinical evaluation with the CAD workstation will be carried out. Then, a second mammography audit will be conducted for the subsequent six-months period without the CAD workstation. We believe that this second segment will be useful to obtain additional baseline data and also to examine the potential variation in the baseline data without the CAD workstation being used clinically. For the final two-year period, the second clinical evaluation of the CAD workstations will be carried out. We will audit the total of three-year periods when the CAD workstations were used and compare those results to the audit of the two six-month audits. This will allow us to study the effects of CAD by comparing parameters such as sensitivity, call-back rates, positive predictive value, etc.

For daily clinical evaluation of the CAD workstations, all screening mammograms will be digitized by a research technologist at each of the two sites and the computed results from the CAD schemes will be stored. When the radiologist reads the original film mammogram, he/she enters his/her findings on normal or abnormal lesion(s) into the CAD workstation using a light pen and soft copy of the mammograms on CRT monitors. Then, the computer output will be indicated on the monitor. The radiologist will then have an opportunity to modify his/her opinion using the light pen. If the radiologist changes his/her initial diagnosis due to the computer output, then the radiologist will enter the final

result into the computer. With this procedure, we will be able to determine the number of breast cancer cases on which the radiologist may miss the lesion initially but may correct his/her findings using the CAD output.

(6) Analysis of outcomes from clinical evaluation of prototype mammography CAD workstations

The effect of mammography CAD workstations on clinical outcomes in the detection of breast cancer will be analyzed both prospectively on a daily basis using the workstation and on a semi-annual basis using the results of mammography audits. Radiologists' performance will be evaluated as a group and also as individuals in order to examine the inter- and intra-observer variability. Since each of the two clinical sites has already established its own mammography audit system, data for "truth" in terms of normal/abnormal (breast cancer) cases will be obtained from each site's mammography audit system for analysis of outcomes in this demonstration project.

2.2 Results and Discussion

(1) Development of automated detection scheme for clustered microcalcifications

We have incorporated local edge-gradient analysis techniques and shift-invariant neural networks for removal of false positives in our CAD scheme for the automated detection of clustered microcalcifications. Therefore, our current CAD scheme contains a large number of parameters such as filter weights, threshold levels, and region of interest (ROI) sizes. The choice of these parameter values determines the overall performance of the system, and thus must be carefully selected. Unfortunately, when the number of parameters becomes large, it is very difficult to obtain the optimal performance, especially when the values of the parameters are correlated with each other. In our preliminary study, we attempted to achieve the optimal overall performance by developing an automated method for determination of the parameter values that maximize the performance

of a mammographic CAD scheme. Our method utilizes a genetic algorithm to search through the possible parameter values, and returns the set of parameters that minimize a cost function which measures the performance of the scheme. Using a database of 89 digitized mammograms, our method demonstrated that the sensitivity of our CAD scheme can be increased from 79% to 87% at a false positive rate of 1.0 per image. The average performance of our CAD scheme on unknown cases is estimated by performing jackknife tests, which was previously not feasible when the parameters of the CAD scheme were determined manually.

(2) Development of automated detection scheme for masses

We have incorporated three techniques to improve the overall performance of our CAD schemes for detection of masses. Three techniques include Hough spectrum analysis, gradient and circularity analysis, and artificial neural networks. We attempted to achieve the high overall performance by optimal selection of many parameters involved in this scheme and also to examine various classifiers to distinguish between lesions and false positives. In our CAD scheme, many features are extracted from potential lesion sites and merged into a single decision variable using a classifier. Numerous features can be extracted from potential lesion sites making it difficult to optimally choose representative features to be used as inputs to a classifier. We have undertaken the problem of feature selection for two different classifiers using a dataset consisting of features extracted from lesions and false-positive detections. We have applied traditional feature selection techniques such as single feature selectors and stepwise selectors. In addition, we have applied genetic algorithms to this search task. A genetic algorithm is an optimization technique loosely based on natural selection. Multiple solutions to a problem are randomly generated and their "fitness" is evaluated. Solutions with better fitness values are more likely to survive to subsequent generations, while solutions with a poor fitness value will "die out." This "survival of the fittest" strategy usually results in a rapid convergence to the

optimal solution. By employing genetic algorithms, we have improved the A_Z of our mass CAD scheme from 0.96 to 0.98 using artificial neural networks. With linear discriminants, the A_Z improved from 0.93 to 0.95. The results from the linear discriminant analysis show that the genetic algorithm feature selection method is as good as, if not better than the stepwise method. Similar results were obtained for the artificial neural network classifiers but the results were not as strong. As with all studies employing neural networks, it is possible that there is over-fitting of the data. We attempted to minimize this effect by simplifying the structure of our networks and by employing cross-validation or leave-one-out tests. Future work will include investigations performed on larger data sets.

(3) Development of automated scheme for characterization of clustered microcalcifications

The automated classification of clustered microcalcifications first segments the individual microcalcifications comprising the cluster from the image. This was achieved by using a combination of background-trend correction with a third order polynomial and grey-level thresholding using 50% of the peak contrast as the threshold value, operating on 99x99-pixel ROIs centered on a microcalcification. From these segmented microcalcifications, features related to individual microcalcifications and the cluster are extracted by the computer. Features of the microcalcifications are: size, irregularity of the microcalcifications, and variation in size and shape of the microcalcifications in a cluster. Features of the distributions of the microcalcifications include the size and circularity of the microcalcification cluster, and the number of microcalcifications. An artificial neural network is used to classify benign versus malignant clusters of microcalcifications using 8 features. The architecture of the ANN is eight input units, six hidden units, and one output unit. The output of the ANN is converted to a likelihood of malignancy.

One-hundred clinical images were obtained from 53 patients biopsied for the suspicion of breast cancer based on clustered microcalcifications. Using ROC analysis, the

computer's A_z value was 0.92 compared to 0.89 for the average of five radiologists. However, in the important clinical region of the ROC curve, which corresponds to the high sensitivity region, the partial A_z was for the computer (0.082) was statistically significantly higher than the value for the radiologists (0.042) with a p-value of 0.03. The partial A_z index measures the area under the ROC curve between 90% and 100% sensitivity and therefore, emphasizes high sensitivity performance compared to A_z , which measures the area under the total curve (ref. 38) The computer scheme correctly identified 82% of the benign patients, all of whom had biopsies (i.e., the radiologist thought the microcalcifications were suspicious for malignancy), and 100% of the malignant patients. On the same set of images the average of five radiologists was only 27% correct in classifying lesions as benign at 100% sensitivity ($P<0.009$).

(4) Development of automated scheme for characterization of masses

The automated classification of masses begins by segmenting the lesions using a grey-level region growing applied to a 512x512 ROI (region of interest) centered on the lesion after background-trend correction (using a second order polynomial) and histogram equalization. The grey-level threshold value is determined from a "transition point." The transition point is the grey level for which there is a discontinuous decrease in the circularity and a corresponding discontinuous increase in size of the grown lesion (ref. 39).

From the segmented lesion, four features related to the degree of spiculation, margin sharpness, density of each mass, and the texture within the mass are extracted automatically from the neighborhoods of mass regions. The techniques for extracting these four features are described in ref.(39). Because of its strong ability to differentiate benign from malignant masses, degree of spiculation is first used in a rule-based technique (i.e., a threshold is applied to the degree of spiculation measure). Those masses that have a spiculation measure lower than a threshold value are then subjected to the ANN, where the remaining features are used as input. The architecture of the ANN is three input units, two

hidden units, and one output unit. The spiculation measure and the output of the ANN are used to determine the likelihood of malignancy.

Using a database of 95 mammograms containing masses from 65 patients (all but one having been biopsied for the suspicion of breast cancer), the performance of the mass classification technique was measured and compared to the results of interpretations by radiologists reading the same cases. Using ROC analysis, the computer classification scheme yielded an A_Z value of 0.94, similar to that of an experienced mammographer ($A_Z=0.91$) and statistically significantly higher than the average performance of the radiologists with less mammographic experience ($A_Z=0.80$). With the database we used, the computer scheme achieved, at 100% sensitivity, a positive predictive value of 83%, which was 12% higher than that of the experienced mammographer and 21% higher than that of the average performance of the less experienced mammographers at a p-value of less than 0.001.

(5) Development of prototype CAD workstation

Our intelligent workstation consists of an IBM RISC 6000 Powerstation 590, a Konica LD4500 film digitizer, an Alphatronix Inspire magneto-optical jukebox, 2 Imlogix 1000 CRT monitors and a Seikosha VP4500 thermal printer. The system has been used in the clinical reading area of the Department of Radiology since November 8, 1994.

Each day all screening mammograms (4-views per case) were digitized. As the films are being digitized, using a 100 micron pixel size and 1024 grey levels, the microcalcification detection program is run on-line in parallel. The mass detection program is run off-line overnight, since the films are not reviewed until the next day. After all four films have been analyzed, the results of the microcalcification detection program are displayed in a single 1024x1280 image as a collage of four 512x620 images with arrow(s) displayed on the image as annotation indicating the computer results. The results were then recorded on thermal paper, upon which the radiologists can make notes and comments.

The results of the mass detection program were printed using the same format the next morning. A full case, four films, can be processed in less than 5 minutes.

Recently, we have made a major modification in the existing workstation by incorporating a touch-screen CRT monitor to display the results of the computer analyses to the radiologist. This will replace the thermal paper copy and will facilitate recording of radiologists' findings. The touch-screen system is used for recording the location of lesions that the radiologist believe are malignant. A digital copy of the four views will be displayed on a monitor with no computer results. The radiologist, after reading the original film mammograms, will touch the screen of the CRT monitor to indicate region(s) in the images that may contain cancer. If the radiologist considers no cancer lesion to be present in the image, he/she will also enter this initial normal finding to the workstation using the touch screen, using a button displayed on the CRT monitor outside the breast region. Once this is done, the computer results will be displayed on the CRT monitor and the radiologist, after reviewing the computer results together with the original films, will again use the touch screen to indicate suspicious region(s) in the images on the monitor, if the location of the malignant region (s) found with the computer output is different from the initial location, or if the initial finding is normal.

(6) Clinical evaluation of CAD workstations

As of October 1997, over 12,000 cases have been analyzed by using our CAD workstation at the University of Chicago. We are analyzing the sensitivity and false-positive rate of the intelligent workstation for the first two-years of implementation: November 8, 1994 to November 7, 1996, which includes 8035 mammographic screening cases. Thirty-five cancers have been confirmed to date within this 2-year period, with one case yielding a negative mammogram but with a palpable lesion. Twenty-three of the 34 cancers were detected by the computer (16 of 23 cases containing masses and 7 of 13 cases containing clustered microcalcifications.) Nine of the patients with cancer had 2 screening

exams during the two-year period. In three of the nine cases, the computer indicated the region in the first exam where the cancer was subsequently diagnosed by the radiologist in the second exam.

The computer output contains, on average, 0.9 false-positive microcalcification clusters and 1.4 false-positive masses. The types of false-positive detections found by the computer in mass detection and clustered microcalcification detection were investigated for 1296 cases. Of the false positives that were indicated by the computer, over 80% of the mass false positives were due to nodular densities on the film. Nearly 50% of the clustered microcalcification false positives were due to arterial calcifications or obviously benign calcifications. In a separate study with clustered microcalcifications, it was found that there is a less than 1% overlap between computer-generated false positives and radiologists-located false positives.

In order to determine the effect of false-positive detections on mammographic interpretation, we calculated the call-back rate in one-year time periods before and after implementation of the workstation in the clinical area. The callback rate is the fraction of screening mammograms read as abnormal. Before introduction of CAD, 13.2% of screeners were called back for further workup and after the introduction of CAD, 12.6% of screeners were called back for further workup. Thus, the false-positive output from the computer did not increase the number of women called back.

Radiologists also subjectively rated the usefulness of having the computer output as an aid in the screening process. The radiologists rated the computer output as beneficial in approximately 5% of the screening cases. It is important to note that in a general screening population, less than 1% of the cases will have cancer.

Concerning the mammography audit without CAD workstation, collection of baseline data on mammographic interpretations performed without computer assistance at Grant Square Imaging, which is a radiology clinic of LaGrange Memorial Hospital, has begun as of January 1997. All mammograms performed at Grant Square were included, as

were mammograms performed at two additional exam sites, Fairview Radiology and Westchester Breast Care Center, which are also read at Grant Square. The examination volume at these sites combined is currently approximately 5,000 studies per year.

Report forms that include basic patient information and mammographic impression have been completed for each patient examination by one of the three study radiologists and retained. Radiologists' recommendations are also categorized on a numerical scale that is similar to the ACR BIRADS scale.

Patient call back rates are being monitored on a monthly basis both per site and per radiologist. These currently run less than 5%. Information on cancer detection rates is being accumulated on a quarterly basis. This is obtained through review of ultrasound and biopsy results from Grant Square Imaging and LaGrange Memorial Hospital. Letters have been sent to physicians whose patients have had an indeterminate or suspicious mammograms but for whom follow-up results are not available. These are generated approximately two months after initial mammographic interpretation. Approximately 80% of patients have had previous recent mammograms, most often at a study site. Current cancer detection rates are similar to expected breast cancer incidence in the population, between 3 and 4 per thousand. The high rate of patient returns to the study sites will allow some assessment of false negative mammography rates to be made. The installation of R2's CAD workstation at LaGrange Memorial Hospital has been delayed due to the cost of transportation and maintenance, which has not been included in the budget, requested by R2. At present we have been negotiating the installation of R2's prototype system in the near future.

2.3 Recommendations in relation to the Statement of Work

Our progress follows closely the proposed statement of work except a delay in the installation of R2's CAD workstation at LaGrange Hospital. However, since we expect

that the CAD workstation will be installed in the near future, we do not recommend a change in the proposed statement of work.

3. CONCLUSIONS

We have made significant progress in the development of various CAD schemes for detection and characterization of breast lesions. Evaluation of our CAD workstation and collection of mammographic audit data have begun. Therefore, it is expected that this project will produce a useful result concerning the impact of CAD schemes in the additional detection of breast cancer.

4. REFERENCES

- (1) Wingo PA, Tong T, Bolden S: Cancer Statistics, 1995. CA 45: 8-30, 1995.
- (2) Lissner J, Kessler M, Anhalt G: Developments in methods for early detection of breast cancer. In: Early Breast Cancer by J. Zander and J. Baltzer, eds. (Springer-Verlag, Berlin 1985), pp. 93-123.
- (3) Bassett LW, Gold RH: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging, (Grune and Stratton, New York, 1987).
- (4) Andersson I, What can we learn from interval carcinomas? Recent Results in Cancer Research 90: 161-163, 1984.
- (5) Martin JE, Moskowitz M, Milbrath JR: Breast cancers missed by mammography. Am. J. Roentgenol. 132: 737, 1979.
- (6) Holland T, Mrvunac M, Hendricks JHCL, Bekker BV: So-called interval cancers of the breast. Pathologic and radiographic analysis. Cancer 49: 2527-2533, 1982.
- (7) Buchanan JR, Spratt JS, Heuser LS: Tumor growth, doubling times, and the inability of the radiologist to diagnosis certain cancers. Radiologic Clinics of North America 21: 115-126, 1983.

- (8) Chan H-P, Doi K, Vyborny CJ, Schmidt RA, Metz CE, et al.: Improvement in radiologists' detection of clustered microcalcifications on mammograms: The potential of computer-aided diagnosis. *Invest Radiol* 25: 1102-1110, 1990.
- (9) Astley S, Hutt I, Adamson S, Miller P, Rose P: Automation in mammography: computer vision and human perception. *Proc. SPIE* 1905: 716-730, 1993.
- (10) Doi K, Giger ML, Nishikawa RM, Hoffmann KR, MacMahon H, Schmidt RA, Chua KG: Digital radiography: A useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images. *Acta Radiologica* 34: 426-439, 1993.
- (11) MacMahon H, Doi K: Digital chest radiography. *Clinics in Chest Medicine* 12: 19-32, 1991.
- (12) Giger ML, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: 3. Automated detection of nodules in peripheral lung field. *Med Phys* 15: 158-166, 1988.
- (13) Matsumoto T, Yoshimura H, Doi K, Giger ML, et al.: Image feature analysis of false positives produced by an automated computerized scheme for the detection of lung nodules in digital chest radiographs. *Invest Radiol* 27: 587-597, 1992.
- (14) Katsuragawa S, Doi K, MacMahon H, et al.: Quantitative analysis of lung texture in the ILO pneumoconioses standard radiographs. *RadioGraphics* 10: 257-269, 1990.
- (15) Chen X, Doi K, Katsuragawa S, MacMahon H: Automated selection of regions of interest for quantitative analysis of lung textures in digital chest radiographs. *Med Phys* 20: 975-982, 1993.
- (16) Nakamori N, Doi K, Sabeti V, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Automated analysis of sizes of heart and lung in digital chest images. *Med Phys* 17: 342-350, 1990.
- (17) Nakamori N, Doi K, MacMahon H, Sasaki Y, Montner S: Effect of heart size parameters computed from digital chest radiographs on detection of cardiomegaly: Potential usefulness for computer-aided diagnosis. *Invest Radiol* 26: 546-550, 1991.

(18) Sanada S, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Automated detection of pneumothorax in chest images. Med Phys 19: 1153-1160, 1992.

(19) Hoffmann KR, Doi K: Correspondence regarding "Determination of instantaneous and average blood flow rates from digital angiograms using distance-density curves." Invest Radiol 27: 274, 1992.

(20) Alperin N, Hoffmann KR, Doi K, Chua KG: Automated analysis of coronary lesions from cineangiograms using vessel tracking and iterative deconvolution techniques. Proc. of Computers in Cardiology, Jerusalem, Israel, Sept. 19-22, IEEE, pp. 153-156, 1990.

(21) Caligiuri P, Giger ML, Favus M, Jia H, Doi K, Dixon LB: Computerized radiographic analysis of osteoporosis. Radiology 186: 471-474, 1993.

(22) Giger ML, Vyborny CJ: CAD in mammography: rationale, methods and possible scenerios. Diagnostic Imaging, June, pgs. 98-113, 1993.

(23) Giger ML: "Future of Breast Imaging. Computer-Aided Diagnosis". Book chapter in: AAPM/RSNA Categorical Course on the Technical Aspects of Breast Imaging, (eds.) Haus A. and Yaffe M. pgs. 257-270, 1992.

(24) Chan HP, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM: Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography. Med Phys 14: 538-548, 1987.

(25) Chan HP, Doi K, Vyborny CJ, Lam KL, Schmidt RA: Computer-aided detection of microcalcifications in mammograms: Methodology and preliminary clinical study. Invest Radiol 23: 664-671, 1988.

(26) Wu Y, Doi K, Giger ML, Nishikawa RM, : Computerized detection of clustered microcalcifications in digital mammograms: Application of artificial neural networks. Med Phys 19: 555-560, 1992.

(27) Zhang W, Itoh K, Tanida J, Ichioka Y: Parallel distributed processing model with local space-invariant interconnections and its optical architecture. Applied Optics 29: 4790-4797, 1990.

(28) Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA: Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. Med Phys 21:517-524, 1994.

(29) Yin FF, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. Med Phys 18: 955-963, 1991.

(30) Yin FF, Giger ML, Vyborny CJ, Doi K, Schmidt RA: Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. Invest Radiol 28: 473-481, 1993.

(31) Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Investigation of feature-analysis techniques. J. Dig. Img 7:18-26, 1994.

(32) Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. Med Phys 21:445-452, 1994.

(33) Nishikawa RM, Giger ML, Doi K, Metz CE, Yin F-F, Vyborny CJ, Schmidt RA: Effect of case selection of the performance of computer-aided detection schemes. Med Phys 21:265-269, 1994.

(34) Nishikawa RM, Haldemann RC, Papaioannou J, Giger ML, Lu P, Schmidt RA, Wolverton DE, Bick U, Doi K: Initial experience with a prototype clinical "intelligent" mammography workstation for computer-aided diagnosis. Proc SPIE 2434: 65-71, 1995.

(35) Fukushima K, Miyake S, Ito T: "Neocognitron: A neural model for a mechanism of visual pattern recognition," IEEE Trans. Systems Man and Cybernetics SMC 13: 826-843, 1983.

(36) Fukushima K: "A neural network for visual pattern recognition," Computer 21: 65-76, 1988.

(37) Schmidt RA, Nishikawa RM, Schreibman K, Giger ML, Doi K, Papaioannou J, Lu P, Stucka J, Birkhahn G: Computer detection of lesions missed by mammography. Proc 2nd Int Workshop on Digital Mammography, Edited by Gale AG, et al, published by Elsevier, pp. 289-294, 1994.

(38) Jiang Y, Metz CE, Nishikawa RM: An ROC partial area index for highly sensitive diagnostic tests. Radiology 201: 745-750, 1996.

(39) Huo Z, Giger ML, Vyborny CJ, et al.: Analysis of spiculation in the computerized classification of mammographic masses. Med Phys 22: 1569-1579, 1995.